

郵送調査における返送バイアス補正の 試みとその妥当性

田 中 富 士 夫

一、問題と目的

近年、社会調査に推計学的手法が適用されてくるに伴い、サンプリング調査の技法は著しく発展してきた。サンプリング調査においては調査（測定）に付随する誤差は、サンプリング誤差と非サンプリング誤差に分けることができる。前者は標本抽出理論によりその大きさを決めコントロールすることができ、誤差の範囲を確率的に表現することが可能である。ところが、後者に属する誤差は統計理論の関知しない誤差であり、サンプリング調査だけに限らず悉皆調査においても回避し得ない種類の誤差である。従って、非サンプリング誤差を如何にして小さくし、排除するかは社会調査の技術における重要な問題である。非サンプリング誤差は、調査のあらゆる段階に介入してくる種々の歪み（偏向）である。この種の偏向（bias）としては、例えば interviewer bias, questionnaire の構成に基づく bias 及び volunteer bias (nonresponse bias) などがあげられる。ここで volunteer bias or nonresponse bias というのは、当初の研究目的のために選ばれた標本集団について調査を行う際、何らかの理由で調査不能となる対象者があるため結局、調査可能対象の結果のみを分析していかねばならないことに基因する偏向効果のことである。従って、面

接調査、留置調査、郵送調査いずれの場合に於いても標本集団から一〇〇%の資料が得られる時にはこのバイアスは問題にはならないが、回収率が一〇〇%を下廻る場合には常に問題になってくる。わが国においても例えば面接調査の場合の調査不能の問題は現在までに若干の研究が行われている(1)二五七―二八〇、(2)三五五―三七二)。

ここでは郵送調査の場合の volunteer bias をとりあげる。それは、郵送調査の回収率(返送率)は面接調査の回収率より遙かに低いことが知られているから、おそらくこのバイアスは郵送調査において著しいと考えられるからである。従来から、volunteer bias の研究には調査における volunteer の社会学的・心理学的特性の追求を目的とする研究が多いが、本論文はむしろ実地的な目的を指摘している。

郵送調査の際、返送された結果だけから全体(発送した全対象者)の結果を推測するための方法、即ち返送資料から非返送者の結果を推定し volunteer bias を補正する方法を提出し、その実地的妥当性を検討するのがこの論文の目的である。

二、補正方法のモデル

問題としてとりあげる調査項目としては、意見、態度、事実等の通常調査されるどのような項目でもよいが、ここでは一応その標識が比率尺度で測定される場合を例にとる。

ある母集団 X_i ($i=1, 2, 3, \dots, N$) の平均 \bar{X} を推定するためには、標本 x_j ($j=1, 2, 3, \dots, n$) の平均 \bar{x} を求めればよいことが知られている。即ち、 \bar{x} は \bar{X} の不偏推定量である。しかし実際の調査では、 n 人の標本は調査可能な n_1 人と調査不能の n_2 人とからなり ($n_1 + n_2 = n$)、 n_2 人については資料が得られない。郵送調査の場合についていえば、 n は調査票を送付した発送総数、 n_1 は返送者数、 n_2 は回答を送り返して呉れない非返送者数である。今、返送対象 n_1 人の平均を \bar{x}_1 、非返送対象 n_2 人の平均を \bar{x}_2 とすると、 \bar{x} は次式で表すことができる。

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \dots \dots \dots (1)$$

この式から \bar{x} を求めるためには \bar{x}_2 が知られねばならないが、その性質上 \bar{x}_2 は未知の値である。従って普通は、 \bar{x}_1 とみなして結果を処理している。若し \bar{x}_1 ならば、標本の大きさが小さくなるため精度は低下するが偏りは生じない。しかし一般には \bar{x}_1 が成立するか否かは不明であるから、むしろ \bar{x}_1 と考えておく方が適當である。そして \bar{x}_1 の時、 \bar{x}_1 の近似式を用いると、結果は或るバイアスを受けることになる。このバイアスを補正するために既知の資料から(1)式の \bar{x}_2 を推定する必要がある。この \bar{x}_2 を推定する方法として幾つかのモデルを考えることができる。

(1) 督促標本法

第一の方法は n_2 人の対象からできるだけ資料を獲得する方法である。通常、郵送調査では返送率が低いから返送しない群に対して幾度か返送依頼の督促を行わねばならない。督促の効果は一般にかなり認められているが、勿論、督促による返送率の増加には限界がある。督促によって返送した n_2' 人 ($n_2' < n_2$) の平均を \bar{x}_2' とすれば、この \bar{x}_2' を \bar{x}_2 の推定値とみなす。即ち、次の近似式(2)を用いて \bar{x} を求める。

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2'}{n_1 + n_2} \dots \dots \dots (2)$$

この方法は、 n_2' 人が n_2 人のランダム・サンプルであるときとみなせる場合には妥当であるが(精度は低くなるが)、この仮定は一般に確かめられていない。この方法を仮りに、督促標本法と名付けることにする。

(2) 遅延標本法

第二の方法は督促によらずに \bar{x}_2 を推定する一つの方法である。郵送調査の返送状況は普通、調査票に指定された締切日迄にかなりの返信が届きその後は相当長い期間に亘って返信がまばらにあり遂に返送が途絶えるという経過をと

ることが知られているから、非返送者は返送迄の期間が非常に長い乃至は無限大の対象者と考えることができる。従って、例えば締切期日以後おくれて返送してくる対象者と非返送対象者は、ある類似した性格を有することが予想されるから、締切期日（あるいは一定の期日）以降におくれて返送してきた n_1' 人（ n_1' は n_1 の一部）の平均 \bar{x}_2' を推定することができるかもしれない。即ち、おくれて返送した対象群を非返送群の標本とみなすわけである。従って次の式(3)によって \bar{x} を求めようとする。この方法を遅延標本法と呼ぶことにする。

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2'}{n_1 + n_2} \dots\dots\dots (3)$$

(3) 時系列法

この方法は第二の方法を更に推し進めた方法である。調査票を發送してから返送されてくる状況を時間的経過の観点から眺めれば、返送状態は時系列をなしていると考えられる。従って若し一定期間毎の返送者についての平均値を時間軸上にプロットした時、ある傾向が認められるとすれば、その傾向が非返送者にもあてはまると考えて、非返送者の平均値を予測することができよう。最も簡単な方法としては時間に基く回帰直線をあてはめる手続をとればよい。今、時間軸上に t_1, t_2, \dots, t_m の m 個の時点を取り夫々の時点における調査項目の平均値を $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ ($\bar{x}_1 = \frac{1}{f} \sum \bar{x}_{1i}$) としこの返送者の資料から例えば最小自乗法で、 $x = a + bt$ なる直線を求める。（但し t は時点、 x はその時点における平均値、 $b \neq 0$ ）

次に非返送群は t_1 以後の時点に引続いて返送されてくると仮定し、 m 個の時点 ($t_{m+1}, t_{m+2}, \dots, t_{m+n}$) をとって夫々の時点で n_2 人づゝ返送されるとすれば、前記の直線式により m 個の x を求め、これを次式に代入して \bar{x} を推定する。この方法を時系列法と呼ぶことにする。

$$\frac{n_1 \bar{x}_1 + \sum_{i=1}^n (a + b f_i) \frac{n_i}{m}}{n_1 + n_2} \dots\dots\dots (4)$$

(4) 標本構成法

この方法は時間的経過には関係なしに \bar{x} を推定する。

従来の諸研究によれば一般に返送群と非返送群は諸種の社会的、心理的特性に差があると報告されている。若し、volunteer bias のため $n_1 \neq n_2$ となるならば、それはある社会的、心理的標識について両群の標本構成が異なることが一因となっているかもしれない。勿論、調査項目に対する反応が社会的・心理的標識と無関係ならばこのようなことは起り得ないし、また \bar{x}_1 と \bar{x}_2 の差がすべてそのような標識の分布の差で説明し尽されるというのではない。今、対象者は凡てある標識について a と b に二分されるものとし、 n 人の対象者が n_a 人と n_b 人づくりに分れ ($n_a + n_b = n$)、

返送群では $n_{1a} + n_{1b} = n_1$ 、非返送群では $n_{2a} + n_{2b} = n_2$ とする。若し $\frac{n_{1a}}{n_1} + \frac{n_{2a}}{n_2} + \frac{n_a}{n}$, $\frac{n_{1b}}{n_1} + \frac{n_{2b}}{n_2} + \frac{n_b}{n}$ で $\bar{x}_a + \bar{x}_b$ (但し \bar{x}_a は n_a 人の平均、 \bar{x}_b は n_b 人の平均) であるとすれば、 \bar{x} の近似値を次の式で求めようとする。即ち、 \bar{x}_{1a} を n_{1a} 人の平均、 \bar{x}_{1b} を n_{1b} 人の平均とすれば、

$$\begin{aligned} & \frac{n_a}{n_1} \frac{\bar{x}_{1a}}{n} + \frac{n_b}{n_1} \frac{\bar{x}_{1b}}{n} \\ & \frac{n_a}{n} \bar{x}_{1a} + \frac{n_b}{n} \bar{x}_{1b} \\ & \text{OR} \\ & \frac{n_1 \bar{x}_1 + n_{2a} \bar{x}_{1a} + n_{2b} \bar{x}_{1b}}{n_1 + n_2} \end{aligned} \left. \dots\dots\dots \right\} (5)$$

しかし、普通、非返送群について知りうる標識は少数のものに限られている。例えばサンプリング台帳に基いて標

本を抽出する時には、台帳の上で知りうる若干の標識（台帳の種類により異なるが、性・年令・居住地域等）しか利用できない。この方法は、返送群の構成を全体の構本構成に合致するよう補正していく方法であるから標本構成法と名付けよう。

以上四種の補正法は必ずしもサンプリング調査を前提としていないから悉皆調査の場合にも適用することができる。要するに、volunteer bias によって生ずる歪みを (3) (4) (5) で表すとすれば、近似式 (2) (3) (4) (5) で行う補正はこの値をなるべく小さくしようとの試みであるといえる。それが実際に歪みを補正し得たか歪みが妥当性である。尙、資料の關係上、本稿では (3) (4) (5) の妥当性を吟味する。

三、方法

一 般 的 方 針

補正方法のモデルとして提出した前記の三方法は、夫々幾つかの仮定に基いて成立している近似法であるから、その妥当性は実証的資料に照合して評価されねばならない。即ち、非返送群の結果の推定値が、真の非返送群の値にどれだけ一致しているかによってその価値が判断される。ところが、返送しない対象の資料なしに真の値を知ることが不可能である。従って、こゝでは次に記すような手続きをとることにする。

調査内容としては対象者の一日における新聞閲読時間を選び、まずこれを面接調査により直接質問して回答をうる。次に面接で回答を得た対象者に対して、その後郵送調査を行い同じく新聞閲読時間の質問を行う。郵送調査の返送を得るにより返送群と非返送群が区分され、返送群の新聞閲読時間を知ることができる。この郵送調査の結果得られた新聞閲読時間と以前に面接で得られた新聞閲読時間とが、ほぼ等しければ、郵送調査で知ろうとした（が知ることのできない）非返送群の新聞閲読時間は面接時のそれと略等しいと仮定できよう。そして、返送群に於いても

郵送調査の結果を利用しないで面接時の結果を用いる。それは同一時期の条件の等しい両群の資料を比較するためである。

具体的手続

最初の面接調査は、ある市場調査の結果を利用した。昭和三十六年七月中旬、石川県下の各市町村に人口に比例して一〇〇地点を割当て、一地点一二名の成人を抽出した。サンプリング台帳は住民登録票を用い、世帯はランダムに抽出したが、対象者は世帯内の成人男女をほぼ同数抽出する意図の下に調査が行われたので厳密な意味ではランダム・サンプリングとはいえない。合計一二〇〇名の対象者について行った面接調査で調査可能な人数は一〇九九名であった。このうち、氏名・年令・職業が不明確な対象を除いたので、次の郵送調査に用いた対象者は九〇九名になった。面接調査においては調査員が各戸に対象者を訪問し、主としてマス・コミュニケーションに対する接触状況、態度等の質問を行ったが、その中に次の質問が含まれている。

Q「あなたが平素新聞を読むのに費やされる時間は毎日何分ぐらいですか」

この質問に対する応答（自由応答で「〇〇分」の如く答えられる）を新聞閲読時間と名付け、以後本稿では凡てこの項目についての結果を分析する。その他の質問項目の資料はこゝでは利用しないが、年令・職業・学歴・生活水準等の基礎項目についての知見は面接調査の際のフェイス・シートの質問で得たものであって本稿でも結果の分析に利用している。

面接調査後約一ヶ月後の同年八月十三日に九〇九名に対して郵送調査を実施した。官製往復ハガキを用い、さきの面接調査と同一の質問を印刷して発送した。質問内容は新聞閲読時間についての項目以外にラジオの聴取時間、現居住地での居住年数、以前の面接調査についての感想等を含み、八月二〇日迄に到着するよう返送を依頼した。また記入者が指定対象と異なる場合を除くために差出人の記名欄を設けた。その結果、返送された回答は九月一日までに五

第1表 郵送調査の
返送状況

差出日付	返送数
8月14日	6
15日	63
16日	94
17日	78
18日	84
19日	69
20日	48
21日	34
22日	18
23日	9
24日	4
25日	5
26日	2
27日	1
28日	1
29日	3
30日	1
31日	2
9月1日	1
消印不明	10
計	533

四、結 果

(1) 面接調査と郵送調査の比較

まず郵送調査の結果と一ヶ月後に行われた面接調査の結果との一致度を調べる。第2表は五三三名の返送群の面接結果と郵送結果による新聞閲読時間の分布を比較した表である。新聞閲読時間は面接では平均三九・二分（標準偏差二八・五分）、郵送では平均三六・九分（標準偏差二四・三分）でも検定によれば両方法の平均値には有意な認められない。また、両方法による新聞閲読時間についてピアソンの相関係数を求めると、 $r=0.98$ となり、これは統計的に充分有意味な相関である。従って、郵送調査の結果と面接調査の結果が等しいことを積極的に主張することはできないにせよ、一方の結果を他方のそれに代用しても著しい誤を犯すことはないであろう。それ故、今後は返送群についても面接調査の結果を利用していく。

(2) 返送群と非返送群の比較

面接調査で得られた新聞閲読時間について返送群と非返送群の比較を行ったのが第3表であ

三八通、他に宛先不明で持戻り四通、また発信宛名（指定対象者）と異なる記名で返送されたもの五通、有効返送数は五三三通であった。従って総返送率は五八・六％である。尚、消印による差出日付別返送状況は第1表に示される。

第2表 同一対象者に実施した面接調査と郵送調査による新聞閲読時間

方法	時間(分) よまぬ	1~10	11~20	21~30	31~40	41~50	51~60	61~70	71~80	81~90	91~	不明	計
面接法		27	61	76	150	164	27	26	2				533
郵送法		10	39	97	192	153	20	14	8				533

第3表 返送群と非返送群の新聞閲読時間の分布の比較(%)

時間 群(分)	よまぬ	1~10	11~20	21~30	31~60	61~90	91~	不明	計
返送群	5.1	11.4	14.3	28.1	30.7	5.1	4.9	0.4	100.0(533)
非返送群	9.6	15.1	18.9	25.8	22.9	3.2	3.4	1.1	100.0(376)
計	6.9	12.9	16.2	27.2	27.5	4.3	4.3	0.7	100.0(909)

る。この表から返送群の方が時間の長い方へ偏っていることが読みとれる。事実、返送群の平均は前記のように三九・二分。非返送群の平均は三一・五分(標準偏差二五・一分)両者の差は1検定で一%水準($t=3.81$)で有意である。つまり、新聞閲読時間についてこの二群は同一母集団からのランダム・サンプルとはいえない。尙、総平均は三六・四分(標準偏差二九・二分)である。この結果から、新聞閲読時間についていえば返送群の資料のみから全体の値を推定すると真の値より長くなることがわかる。

(3) 遅延標本法の妥当性

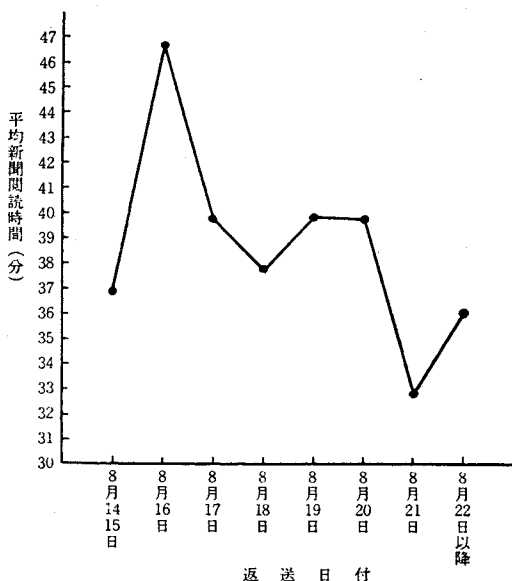
この方法の基礎は返送群のうち、締切期日以降におくれて返送してきた対象はそれ以前に送付した対象に比べて非返送群に近い性格を有しているだろうから、これを非返送群の標本とみなしうるといふ仮定である。

調査票には八月二〇日迄に返送するよう依頼してあったにも拘らず第1表のように実際には二一日以降の消印のものが八十一通あった。今、仮りにこの八一名の新聞閲読時間を以って非返送群のそれに代用することにする。八一名の遅延群では平均値は三四・八分(標準偏差二三・六分)であった。この値を近似式(3)に代入すれば求めるものは三七・四分となり、真のものは三六・四分に近づくことがわかる。

(4) 時系列法の妥当性

第1図は横軸に返送ハガキの差出日付をとり縦軸に各日付毎に算出した新聞閲読時間の平均値を表したグラフである。こゝでは平均新聞閲読時間の差出日による変動が単なる誤差ではなく、或る傾向が存在し時間の函数であると仮定する。かゝる仮定にたてば非返送群つまり未だ回答が到着していない群は返送群の結果から

第1図 返送日付による平均新聞閲読時間の変動



日毎の区分に対応させるためである。従って非返送群は同様に六五人づつのブロックに区切っていけば返送順位第九位から第十四位まで(但し最後のブロックは五一名)のブロックに該当することになる。

次にこのグラフで八個の実測値に対して時間(順序)に基いた回帰線を求める。その一つの方法は最小自乗法によって直線をあてはめることである。第2図の直観的視察からこの時系列的経過が曲線的変動をしているのではないかと疑問が生ずるが曲線式をとること自体には異存がなくてもその曲線型を二次式と想定すべきかあるいは他の函数式と考えるべきかは明らかではない。筆者はこのグラフの経過をある偶然変動を繰返し乍ら次第に下降する傾向とし

外挿法によって予測することができよう。しかしながら第1図の如き表示法では各日付毎の度数が一定でなく(第1表参照)、また時間軸の目盛は等間隔ではないから非返送群の時間的位置づけを行うことが困難である。それ故、こゝでは時間という変数を返送ハガキを差出した順序とみなす。このように考えれば、非返送群は二二日以降の返送者の次に位置づけることが可能である。第2図はこの観点にたつて横軸には回答ハガキの差出順に六五名づつのブロックにまとめた場合の順位をとり、新聞閲読時間の平均値を縦軸にとって時間的経過に伴う変動を示したグラフである。各ブロック六五人としたのは便宜的な手段であつて、各ブロックの度数を一定(但し第八位のブロックは六八名)にし且つほぼ第1図の一

て把え、むしろ直線をあてはめた方が適當と考えた。

今、八個の実測値に基き最小自乗法を用いて回帰直線を求めると次式が得られた。

$$x = 42.95 - 0.75t$$

(但し x はある順位のブロックにおける平均新聞
閲読時間 (分))

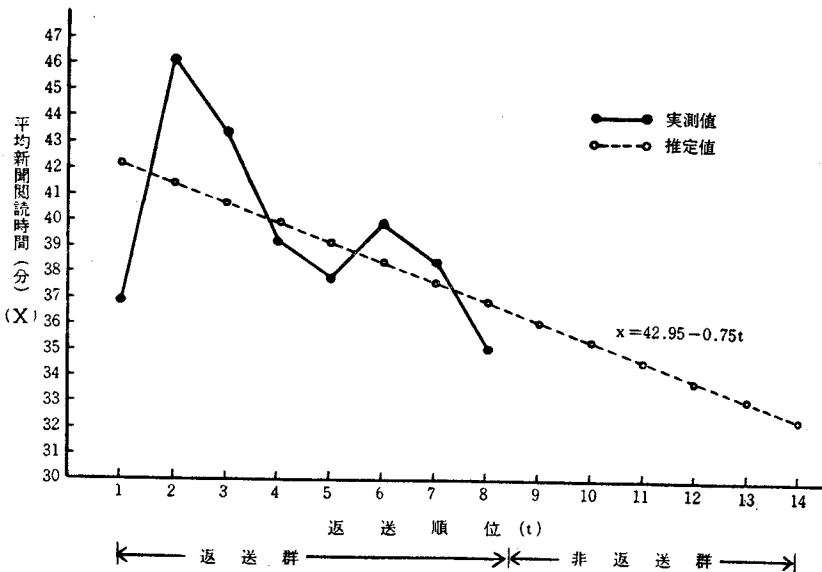
t は返送順序によるブロック順位)

この式によって示される直線は第2図の点線である。次にこの式を用いて非返送群の平均新聞閲読時間を予測した結果を近似式(4)に代入すれば、 \bar{x} は三七・二分となる。この値は \bar{x}_1 よりも真の値 \bar{x} に近づくことになるから時系列法によってバイアスを或る程度補正することが出来る。

(5) 標本構成法

この方法は、返送群の結果を直ちに非返送群の結果の推定値とせずに、ある標識について返送群、非返送群を下位分類しその同一標識をもつ下位群については返送群の値を以て非返送群の値に代用するのである。この方法によってバイアスが補正されると考える

第2図 返送順位による平均新聞閲読時間の変動
と最小自乗法による直線のあてはめ



ためには次の仮定が必要である。

- (a) ある標識（例えば男女）について返送群と非返送群の対象者の構成に差異があること。
 - (b) その標識は問題としている調査項目（今の場合は新聞閲読時間）と無関係ではないこと。
 - (c) 同一標識の下では返送群と非返送群は調査項目に対する反応に差がないこと。
- 補正モデルを吟味するためにはこれらの仮定が成立しうるかを検討することが必要である。

(a) 返送群と非返送群の対象者の構成

本研究の場合は非返送群についても面接調査の結果が利用できるから、一応標識として、居住地域・性・年令・職業・学歴・生活水準を選び両群の対象者の構成を比較したのが第4表である。地域は金沢市・小都市及び郡部に三分した。小都市とは、七尾・加賀・小松・羽咋・珠洲の各市を指し、郡部とはその他の町村である。また生活水準の評定は五種類の所謂耐久消費財（ミシン、洋服ダンス、電気洗濯機、電話、電気冷蔵庫）の所有点数を以て表した。この方法の妥当性と信頼性については既に報告したところである(3)。

ある標識についての両群の差ということは換言すればその標識による返送率の差である。第4表によれば、学歴により返送率に差が認められる外は他の標識は凡て両群間で有意な差が認められない。但し学歴の差は三種の学歴相互間に凡て認められるのではなくて、大学卒・小学卒と中学卒との間にだけ認められる。また差の認められなかった標識についても例えば年令についていえば、三十一才〜四〇才及び六一才以上の年代は共に返送率が低くこれは他の年代とは有意な差($\chi^2=6.7503, p<0.01$)であり、職業分類についていえば農林漁業より主婦の方が遙かに返送率が高く($\chi^2=27.839, p<0.01$)、生活水準では「0」の段階の対象者が「1」以上の対象者より返送率が低いこと($\chi^2=6.3079, p<0.02$)も有意である。しかし返送率における性差や地域差が統計的有意水準に達しないからといって直ちに両群における男女の構成や地域的構成が同一であるといえないのは勿論である。

第4表 返送群及び非返送群の比較

a) 地 域

	金沢市	小都市	郡部	計
返 送 群	194	164	175	533
非返送群	127	102	147	376
全	321	266	322	909

$$\chi^2=3.9003 \quad p>0.10$$

b) 性

	男	女	計
返 送 群	261	272	533
非返送群	176	200	376
全	437	472	909

$$\chi^2=0.0996 \quad p>0.95$$

c) 年 令

	21~30	31~40	41~50	51~60	61~	計
返 送 群	78	145	144	114	52	533
非返送群	51	129	86	67	43	376
全	129	274	230	181	95	909

$$\chi^2=7.1975 \quad p>0.10$$

d) 職 業

	勤人	労務	商工 自営	管 理 自由業	農林 漁業	主婦	無職	計
返 送 群	53	85	63	23	114	170	25	533
非返送群	38	55	47	17	96	105	18	376
全	91	140	110	40	210	275	43	909

$$\chi^2=3.1817 \quad p>0.70$$

e) 学 歴

	新旧大学 高専卒	旧 新 高 卒	中 卒	旧 新 中 卒	小 卒	計
返 送 群	28	160		345		533
非返送群	25	78		273		376
全	53	238		618		909

$$\chi^2=9.9999 \quad p<0.01$$

f) 生活水準*

	0	1	2	3	4	5	計
返 送 群	83	135	118	106	58	31	531
非返送群	83	81	78	67	46	19	374
全	166	216	196	173	104	50	905

$$\chi^2=7.7455 \quad p<0.10$$

*両群に2名づゝ評定結果不明の対象者があつたので表では省かれて
いる。

更に検討すべきことは単一の標識については二群に差がなくても、それらの標識を組合せた場合にその群構成に差が生ずるのではないかという点である。さきにあげた標識のうち非返送群について実際の組合せについて返送率をみると第5表
リング台帳の段階では、三標識（地域・性・年令）だけであるから、これらの組合せについて返送率をみると第5表

のようになる。例えば小都市では概して男より女が返送率高く郡部では反対に男が高い傾向のあることがわかるが共に有意差なく、このように三標識を用いて細分化を行うと一群の例数が少くなりかえって一定の傾向が認められなくなる。少くとも、地域・性・年令については単一の標識のみでは差のなかったものがそれらの組合せによって差が生じて来るといふ場合はないように思われる。

(b) 諸種の標識と新聞閲読時間との関係

返送群と非返送群に或る標識について差があったとしても、新聞閲読時間とその標識とが全く無関係ならば補正はできぬわけである。次に返送群及び非返送群を含めた発送者全体について新聞閲読時間と地域・性・年令との関係を調べる。

地域については、平均新聞閲読時間は、金沢市三八・三分（標準偏差二八・四分）、小都市三八・一分（標準偏差二七・六分）、郡部三三・二分（標準偏差二九・六分）でも検定によれば金沢市と郡部及び小都市と郡部間には共に五%水準で有意な差が認められる。

性については、男四三・八分（標準偏差三一・七分）、女二九・六分（標準偏差二三・五分）で両者の差 ($t=7.394$) は1%水準で有意である。

年令については第6表のように年令の増加と共に平均新聞閲読時間が長くなる傾向が認められ相隣れる年代間には有意差はないがその他の年代間には概して統計的に有意な差が認められる。

以上の結果から標本構成法の第二の仮説は一応承認できる。

(c) 同一標識についての返送群、非返送群の新聞閲読時間の比較

第5表 地域・性・年令別にみた返送率(%)

年令	地域・性		金 沢 市		小 都 市		郡 部	
	性		男	女	男	女	男	女
21~30			100.0	48.0	60.0	68.0	75.0	65.8
31~40			57.8	55.5	46.4	62.9	50.0	42.9
41~50			70.0	71.5	51.5	61.3	59.6	56.3
51~60			54.1	81.2	68.9	77.8	68.4	37.0
61~			56.5	50.0	61.5	33.3	54.9	37.5
計			60.8	60.2	58.5	64.9	59.8	49.4

第6表 年令別平均新聞閲読時間及び年令差の有意水準
(*t*検定)

年 令 (才)		21~30	31~40	41~50	51~60	61~
新聞 閲 読 時 間 (分)	M	29.7	33.4	36.9	40.9	44.8
	σ	23.1	25.2	25.3	34.5	42.4
年 令 (才)	21~30		n. s.	$p < 0.01$	$p < 0.01$	$p < 0.01$
	31~40			n. s.	$p < 0.05$	$p < 0.05$
	41~50				n. s.	n. s.
	51~60					n. s.
	61~					

n. s. : non significant

組は返送群が長く、六組は非返送群が長く、一組は全く等しい値をとっていることがわかる。この結果から返送群が非返送群より平均新聞閲読時間が長いことは符号検定によって1%の有意水準で確かである。つまり地域・性・年令が同一な対象者にあっても返送群は非返送群に比し常に平均新聞閲読時間が長い傾向がある。従って第三の仮定は事実とは一致しないといえる。しかし第一、第二の仮定が認められる限り、標成構成法のモデルによる補正が全く意味がないとはいえない。というのは、第一、第二の仮定は標本構成法で補正するための必要条件でありこの条件が満た

例えば地域・性・年令・職業・学歴・生活水準等の標識が同じ対象者は返送群においても非返送群に於いても平均新聞閲読時間が等しいという仮定がなければ標本構成法のモデルによってバイアスを除くことはできない。勿論、平均新聞閲読時間が前記のような標識だけによって完全に規定されてしまふと考えることはできないから、同一標識をもつ対象者が両群で完全に一致するとの仮定は必要ではないが、少くとも両者が全くかけ離れていたり、一定方向の差が常に認められるようでは、つまり他の標識が同一であっても「返送」、「非返送」という標識によって新聞閲読時間が著しく異なるようでは標本構成法によるバイアスの補正は余り期待できない。

地域・性・年令の三標識が等しい対象者を返送群、非返送群に分けて両群の平均新聞閲読時間を求めたのが第7表である。返送群と非返送群の平均値の差は検定していないが、夫々の例数が少く分散がかなり大であるから多分有意差の認められる場合は少いであろう。しかし夫々同一標識の返送群と非返送群との平均新聞閲読時間について、その長短を比較すると二九組中二二

第8表 標本構成法による平均
新聞閲読時間の補正

補正に用いた標識	補正值(分)
性	39.109
地域	39.117
年令	39.085
性・地域	38.902
性・地域・年令	38.648

れていなければ補正は不可能である。これに對して第三の仮定は、いわば充分条件ともいふべきであつて、この条件が満されていれば多くの標識を組合せて用いれば標本構成法によつて究極的には眞の

が推定できると考えられる。若しこの条件が満されていない場合にはバイアスの補正は或る程度可能であるが、標識数を増加していつても究極において眞のには近づかないで、(81+8)に近づくと考えられる。こゝで α は「返送」、

「非返送」に關係する差である。従つて補正が全く無効ではないがその効果は著しく限定されよう。

次に、近似式(5)により實際に補正を行つてみると、第8表のようになり、いづれの場合も α の値三九・二分から眞の α の値三六・四分に幾分近づくことがわかる。また一個の標識だけを用いるより幾つかの標識を組合せる方が有効なようである。しかしこれらの結果は、補正しない以前の値と比べて大きな差は認められない。このことは、例えば面接調査の調査不能サンプルをこの種の方法で補正した結果(2)、三六五頁、但し、そこで用いられた方法は本稿の補正法と全く同じではない)とよく似ている。

第7表 地域・性・年令別平均
新聞閲読時間(分)

地域			群	返 送 群	非 返 送 群
性	年 令				
金 沢 市	男	21~30	31.7	—*	
		31~40	44.8	43.9	
		41~50	49.1	36.7	
		51~60	54.5	38.2	
		61~	34.6	57.0	
	女	21~30	28.3	26.9	
		31~40	38.8	24.2	
		41~50	35.4	28.2	
		51~60	34.2	30.0	
		61~	70.0	35.0	
小 都 市	男	21~30	33.3	50.0	
		31~40	35.4	32.9	
		41~50	43.8	33.0	
		51~60	52.1	50.0	
		61~	52.7	40.0	
	女	21~30	27.1	29.4	
		31~40	32.6	31.0	
		41~50	46.3	26.7	
		51~60	29.6	17.5	
		61~	30.0	20.0	
郡 部	男	21~30	35.8	15.0	
		31~40	40.7	27.0	
		41~50	39.5	39.7	
		51~60	53.8	33.4	
		61~	46.5	60.8	
	女	21~30	28.2	31.9	
		31~40	26.7	26.7	
		41~50	25.6	22.9	
		51~60	30.5	10.3	
		61~	11.7	7.0	

*該当者なし

五、考 察

郵送調査における volunteer bias 補正の方法として提出された三種のモデルの妥当性を実際の資料に照し合せて検討することがこの論文の目的であった。その結果は、三種のモデルは夫々バイアスを或る程度補正するのに有効であることがわかった。三者の妥当性を比較すると、標本構成法は実際の適用に際してかなりの労力が必要とする割にその効果はそれ程現われない。時系列法は最も有効に思われる。計算に要する労力・時間に比して有効なのは遅延標本法であるといえよう。(更に標識毎に時系列法を行うという両方法の組合せを考えることができるが実用には余りにも煩瑣であろう。)

しかしながら、これらの結果から直ちにバイアス補正についての一般的な結論を導き出すことはできないであろう。というのは、補正の妥当性は調査項目の内容によって、あるいは返送率の高さによって当然変わってくるが予想されるからである。次に夫々の方法について考察する。

時系列法では差出(返送)日付の確認とその取扱いが先ず問題である。本研究の場合でも一〇通の回答が消印なしに到着したため処理できなかった。また対象者に調査票が到着してから返送するまでの期日を変数にしている以上、発送日を適当に調節しないと地域による差を除くことができない。その意味では時間の単位として余りに細分化するのは適当ではない。

対象を返送日付の時間的順序に配列するということは、別の観点からみれば対象者を返送確率の連続体(response-nontresponse probability continuum)の上に位置づけることである。つまり対象者の返送確率(返送する可能性)の大きさを決めるために返送順位という時間的順序を用いていると考えるのである。このように考えると時系列の問題というよりは対象者の返送確率という変数(標識)と調査項目(こゝでは新聞閲読時間)の反応との関係から非返送

群の結果を予測することになる。返送確率については次のように考える。返送群つまり顕現的な返送行動をとる対象は返送確率がある値以上の者であり、その値以下の返送確率の者が非返送群を構成している。そして夫々の群内の対象者の返送確率は一定ではなく変動が存在する。返送確率の大きさは返送群にあっては調査票発送から返送までの期間という測値で表すことができる。非返送群については、若し督促を行った場合に返送してくるかどうか、あるいはその返送までの期間という測値で表わされるであろう。非返送群の返送確率についてのこのような仮定は、既に Zimmer, H. (4) により若干吟味されている。更に何回かの督促を行っても返送しない対象に改めて面接調査を実施すると仮定する。従来から、面接法による回収率は郵送法の返送率よりかなり高いことが知られているから、若し両方法に於ける調査不能対象が異質的でないと仮定すれば（この仮定は必ずしも凡ての場合にあてはまらないが）、面接調査の実施可能性によって更に返送確率（むしろ反応確率というべきかもしれない）の低い段階の対象者を弁別できるかもしれない。本研究では督促をしなかったし、また既に面接不能者を郵送対象に含めなかったので残念乍らの点の吟味を行うことはできない。

時系列に於ける予測法自体の問題は更に検討すべき余地があろう。こゝでは直線をあてはめること自体にもかなり疑問が残されている。しかし変動傾向の型が判っきりしていない時には一般に余りにも複雑な函数式を想定することがそれ程有効であるとは思われない。

遅延標本法の仮定は返送確率の概念でいえば、締切期限後の遅延返送者の返送確率が非返送群の返送確率に近いと考え、返送確率について遅延群は非返送群と同じ母集団に属すると考えていることになる。しかし、遅延群が非返送群と同質と見做し得るか否かは一つには遅延群の定義即ちどの点から遅延群と分割するかの限界と関係が深いであろう。この適切な区切りの決定は一般にはわかっていない。不適切な区分や遅延標本が小さいために、バイアス補正がかえって誤った結論を導き出すかもしれないという危険性はこの種の方法では除くことができない。

標本構成法はその仮定のなかに時間的経過の要因が全く考慮されていない点で他の方法と異っている。この方法は結果のところに記したように三つの仮定を前提としており、調査項目に対する反応が返送・非返送という標識と無相関であることを暗黙のうちに想定しているわけである。しかし、一般に郵送調査ではこの仮定が承認されないように思われる。というのは調査項目に対する対象者の関心は返送率の規定要因として見逃すことができないからである。そして調査項目に対する関心とその項目に対する反応とが関係を有し、調査項目に対する関心と返送行動(volunteering behavior)とに関連が認められるとすれば、項目に対する反応と返送行動が何らかの関係を有することは当然予想されるところである。また一方、若し返送行動が項目に対する関心や反応と無関係に全くランダムに生ずると考えれば、おそらく返送群と非返送群は種々の標識についてはと同様な比率で構成されるだろうからやはり標本構成法による補正は有効とは思われない。もっとも項目に対する関心、反応が返送行動とは無関係だがその他の標識とは関係があるという可能性がないわけではない。若し返送行動がランダムに生ずるとすれば、常に返送するような対象者あるいは常に返送しない対象者の特徴というものは存在しないことになる。調査の目的、内容、形式等の如何に拘らず如何なる場合にも常に終始一貫して返送する対象者と返送しない対象者とがあるとは仮定しないにしても返送行動が全くランダムに生起するといえないことは過去に於ける返送率規定要因の諸研究が示している通りである。

従って返送・非返送によって項目に対する反応に差があるとすれば、前記のように標本構成法による補正にはかなりの限界があるといわねばならない。しかし、限界があることは補正の効果が薄いという結果をうむかもしれないが、少くとも返送群、非返送群の対象者構成の差に基づくバイアスは除きうることを意味する。けれども、返送・非返送に本質的なバイアスは除くことができない。之に反し時系列法、遅延標本法は共に既知の返送群の結果から全く未知の非返送群の結果を予測する方式をとっているのであるから、両者の方法はかなり異った性格をもつわけである。

六、要 約

郵送調査から得られる資料の volunteer bias を補正するためのモデルを提出し、その妥当性を実際の資料に照合して吟味することがこの研究の目的であった。

補正モデルとして(1)遅延標本法、(2)時系列法、(3)標本構成法を提出した。

この三方法による補正の適否を検討するため、予め面接調査によって毎日の平均的な「新聞閲読時間」がわかっている対象者九〇九名に対して、面接後約一ヶ月の後に往復ヘガキにより「新聞閲読時間」を質問した。その結果は九〇九名中五三三名の返送があった。返送群五三三名については面接法及び郵送法によって得られた新聞閲読時間に有意差なく且つ、両者間に有意な相関が認められたので面接の結果を郵送結果と見做し、三種の補正法による非返送群の推定値を実際の非返送群の値と照合した。

結果は、いずれの方法も補正の効果が認められたが、妥当性の程度は、時系列法、遅延標本法、標本構成法の順であった。このうち、手続が簡単な割に有効なのは遅延標本法であった。

三種の補正法はいずれも夫々幾つかの仮定に基いているので、それらの前提となっている仮定について若干の考察を加えた。補正の適切性乃至妥当性は調査内容、返送率その他の条件により夫々変化すると思われるので一般的な結論をこの結果だけから導くことは危険であろう。今後更にこの種の妥当性研究を積み重ねていく必要があらう。

七、文 献

- (1) 林知己夫…サンプリング調査はどう行つか、一九五一、東京、東大出版部
- (2) 統計数理研究所国民性調査委員会…日本人の国民性、一九六一、東京、至誠堂

- (3) 田中富士夫…社会調査における生活水準測定尺度作製の試み、金沢大学法文学部論集、哲学史学篇8、一七七—一九六、一九六〇
- (4) Zimmer, H.: Validity of Extrapolating Nonresponse Bias from Mail Questionnaire Follow-ups. J. appl. Psychol., 1956, 40, 117~121